

2005 Special Issue

Dynamical systems and cognitive linguistics: toward an active morphodynamical semantics

René Doursat^{a,*}, Jean Petitot^b

^aGoodman Brain Computation Lab, University of Nevada, Reno, Reno, NV 89557, USA

^bCREA, Ecole Polytechnique, 1, rue Descartes, 75005 Paris, France

Abstract

We propose a novel dynamical system approach to cognitive linguistics based on cellular automata and spiking neural networks¹. How can the same relationship ‘in’ apply to containers as different as ‘box’, ‘tree’ or ‘bowl’? Our objective is to categorize the infinite diversity of schematic visual scenes into a small set of grammatical elements and elucidate the *topology of language*. Gestalt-inspired semantic studies have shown that spatial prepositions such as ‘in’ or ‘above’ are neutral toward the shape and size of objects. We suggest that this invariance can be explained by introducing *morphodynamical transforms*, which erase image details and create virtual structures or *singularities* (boundaries, skeleton), and call this paradigm ‘active semantics’. Singularities arise from a large-scale lattice of coupled excitable units exhibiting spatiotemporal pattern formation, in particular *traveling waves*. This work addresses the crucial cognitive mechanisms of spatial *schematization* and *categorization* at the interface between vision and language and anchors them to expansion processes such as activity diffusion or wave propagation.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: Cognitive linguistics; Semantics; Spatial cognition; Categorization; Scene recognition; Machine vision; Dynamical systems; Morphodynamics; Cellular automata; Spiking neural networks; Coupled oscillators; Synchronization; Traveling waves; Spatiotemporal patterns

1. Introduction

How can the same English relationship ‘in’ apply to scenes as different as ‘the shoe in the box’ (small, hollow, closed volume), ‘the bird in the tree’ (large, dense, open volume) or ‘the fruit in the bowl’ (curved surface)? What is the common ‘across’ invariant behind ‘he swam across the lake’ (smooth trajectory, irregular shape) and ‘the fly zigzagged across the hall’ (jagged trajectory, regular volume)? How can language, especially its spatial elements, be so insensitive to wide topological and morphological differences among visual percepts? In short, how does language drastically simplify information and *categorize*?

This study examines the link between the spatial structure of visual scenes and their linguistic descriptions. We propose a spiking neural model aimed at mapping the endless diversity of schematic visual scenes to a small set of standard semantic labels (Doursat & Petitot, 2005). Our suggestion is that spatial and grammatical structures are related to each other through image transformations that rely on diffusion processes and wavelike neural activity.

1.1. Toward a protosemantics

Contemporary theories of semantics have shown that terms with spatiotemporal content are highly polysemous. For example, different uses of ‘in’ relate to qualitatively different spatial situations (Herskovits, 1986): ‘the cat in₁ the house’, ‘the bird in₂ the tree’, ‘the flower in₃ the vase’, ‘the crack in₄ the vase’, etc. Spatial prepositions typically form prototype-based radial categories (Rosch, 1975). Here, ‘in₁’ represents the most prototypical containment schema, while ‘in₂’ departs from it slightly in the texture and openness of the container. Top-down, knowledge-based inferences also create *metonymic* effects such as ‘in₃’—the stem, not the flower, is inside the vase. In the subcategorical

* Corresponding author. Tel.: +1 775 784 6974; fax: +1 509 471 3511.
E-mail addresses: doursat@unr.edu (R. Doursat), petitot@polytechnique.fr (J. Petitot).

URLs: www.cse.unr.edu/~doursat (R. Doursat), www.crea.polytechnique.fr/jeanpetitot/home.html (J. Petitot).

¹ An abbreviated version of some portions of this article appeared in Doursat & Petitot (2005), published under the IEEE copyright.

island ‘in₄’, the container is a surface, etc. Another major departure from prototypical use comes from *metaphors*, a pervasive mechanism of cognitive organization rooted in the fundamental domains of space and time (Lakoff, 1987; Talmy, 2000). The schemas grammaticalized by ‘in’ are not just used to structure physical scenes but also abstract situations: one can be ‘in_A a committee’, ‘in_B doubt’, etc. These uses of ‘in’ pertain to a virtual concept of space generalized from real space: ‘in_A’ metaphorically maps to an element of a discrete numerable set, such as ‘in₅ a crowd’, while ‘in_B’ relates to an immersion into a continuous ambient substance, such as ‘in₆ water’.

With these preliminary remarks in mind, we restrict the scope of the present study to relatively homogeneous subcategories or *protosemantic* features, i.e. low-level compared to linguistic categories but high-level compared to local visual features. Our goal is to categorize scenes into elementary semantic subclasses, such as ‘in₁’, or clusters of closely related subclasses, such as ‘in₁–in₂’. We will not attempt to delineate the whole cultural complex formed by the English preposition ‘in’ with a single (and nonexistent) universal. We also leave aside issues of top-down control or ‘intentionality’ deciding how to choose and compose elements. Yet, even in their most typical and concrete instantiation, we are still facing the great difficulty of linking these invariant semantic kernels to the infinite continuum of perceptual shape diversity. To help solve this problem, we focus on a collection of low-level dynamical routines that structure visual scenes in a ‘Gestaltist’ fashion.

1.2. The Gestaltist conception of relations

Through a shift of paradigm initiated in the 1970s, the formalist view that language is functionally autonomous was revised by a set of works (e.g. Lakoff, 1987; Langacker, 1987; Talmy, 2000) collectively named *cognitive linguistics*, for which language is much rather ‘embodied’ in perception, action and inner conceptual representations. In this paradigm, generative grammar models have been superseded by studies of the *interdependence of language and perceptual reality* and how these two systems influence each other’s organization. Cognitive linguistics is directly preoccupied with meaning and categorization. Refuting the distinction between syntax and semantics, it postulates a *conceptual level* of representation, where language, perception and action become compatible (Jackendoff, 1983). It also remarkably revived the Gestalt approach calling into question the traditional roles assigned to visual perception—as a faculty only dealing with object shapes—and language—as a faculty only dealing with relations between objects. Classical formal linguistics follows the logical atomism of set theory: ‘things’ are already individuated symbols and ‘relations’ are abstract links connecting these symbols. By contrast, in the Gestaltist or *mereological* conception, things and relations constitute wholes: relations

are not taken for granted but emerge together with the objects through segmentation and transformation. Here, objects involved in relations are parts of a higher-order complex object—the global configuration.

1.3. Toward an active morphodynamical semantics

A fundamental thesis of Talmy’s Gestalt semantics is that linguistic elements structure the conceptual material in the same way visual elements structure the perceptual material and that this structuration is mostly of *grammatical* origin, as opposed to lexical. Grammatical elements (‘in’, ‘above’) provide the conceptual structure or ‘scaffolding’ of a cognitive representation, while lexical elements (‘bird’, ‘box’) provide its conceptual contents. Consider a figure or *trajector* (TR)—to use Langacker’s terminology—profiled against a ground or *landmark* (LM), for example ‘TR in LM.’ One of Talmy’s most compelling ideas is that LM is not just a passive frame but is actively structured (scaffolded) in a geometrical and morphological way radically different from symbolic relations. In this Gestaltist view, for example, ‘I am in/on/far from the street’ construes LM (the street) either as a volume, a surface or a reference point, thus dividing (scaffolding) reality in conceptually different ways. By contrast, formal theories of syntax assimilate the street to a symbol, irrespective of its perceptual properties, and encapsulate all spatial meaning in an abstract relationship labeled ‘in’.

Now, by giving a new significance to the object’s geometry we are challenged to find how it varies with linguistic circumstances, i.e., how prepositions actually select and process certain morphological features from the perceptual data and ignore others. We propose that simple spatial elements like ‘in’, ‘above’, ‘across’, etc., correspond in fact to visual processing algorithms that take perceptual shapes in input, transform them in a specific way, and deliver a semantic schema in output. We call these algorithms *morphodynamical routines* and the global process *active semantics*. Morphodynamical routines erase details and create new forms that evolve temporally. They enrich a visual scene with virtual structures or *singularities* (e.g., influence zones, fictive boundaries, skeletons, intersection points, force fields, etc.) that were not originally part of it but are ultimately revealing of its conveyed meaning. We suggest below that these routines rest upon object-centered and diffusion-like nonlinear dynamics.

In the remainder of this article, Section 2 applies the active semantics approach to the analysis of spatial grammatical elements and draws a link with visual processing routines in 2-D cellular automata (CA). In Section 3 we discuss spatial invariance and the notion of ‘linguistic topology’ in the context of mathematical geometry, which leads us to reaffirm the importance of transformation routines. Section 4 introduces waveform dynamics in spiking neural networks and shows its role in active semantics as a plausible mechanism of

expansion-based morphology. This is further developed in Section 5 through two models of spatial scene categorization based on wave interaction and detection. We conclude in Section 6 by hinting at future work beyond the preliminary results presented here and pointing out the originality of our proposal compared to both current linguistic and neural modeling.

2. Morphodynamical cellular automata

Numerous examples collected by Talmy show that grammatical elements are largely indifferent to the morphological details of objects or trajectories. Such invariance of meaning points to the existence of underlying visual-semantic routines that perform a *drastic but targeted simplification* of the geometrical data. We examine a representative sample that helps us introduce various types of data-reduction transforms essential to scene categorization.

2.1. Magnitude invariance \rightarrow multiscale analysis

Grammatical elements are fundamentally scale-invariant. The sentences ‘this speck is smaller than that speck’ and ‘this planet is smaller than that planet’ (Talmy, 2000, p. 25) show that real distance and size do not play a role in the partitioning of space carried out by the deictics ‘this’ and ‘that’. The relative notions of ‘proximity’ and ‘remoteness’ conveyed by these elements are the same, whether speaking of millimeters or parsecs. It does not mean, however, that grammatical elements are insensitive to metric aspects but rather that they can uniformly handle similar metric configurations at vastly different *scales*. Multiscale processing therefore lies at the core of Gestalt semantics and, interestingly, has also become a main area of research in computational vision (since Mallat, 1989).

2.2. Bulk invariance \rightarrow skeleton

Another form of magnitude invariance appears in the selective rescaling of some dimensions and not others. This is the case of thickness or bulk invariance. For example, ‘the caterpillar crawled up along the filament/flagpole/redwood tree’ (Talmy, 2000, p. 31) shows that ‘along’ is indifferent to the girth of LM and focuses only on its main axis or *skeleton*, parallel to the direction of TR’s trajectory. Like multiscale analysis, skeleton transforms are widely used in machine vision and implemented using various algorithms. Studied by Blum (1973) under the name ‘medial axis transform’ and by others as ‘cut locus’, ‘stick figures’, ‘shock graphs’ or ‘Voronoi diagrams’ (e.g. Marr, 1982; Siddiqi, Shokoufandeh, Dickinson, & Zucker, 1999; Zhu & Yuille, 1996), morphological symmetry plays a crucial role in theories of perception and is even considered a fundamental structuring principle of cognition

(Leyton, 1992). Skeletons indeed conveniently simplify and schematize shapes by getting rid of unnecessary details, while at the same time conserving their most important structural features. There is also experimental evidence that the visual system effectively constructs the symmetry axis of shapes (Kimia, 2003; Lee, 2003).

2.3. Continuity invariance \rightarrow expansion

Sentences such as ‘the ball/fruit/bird is in the box/bowl/cage’ are good examples of the neutrality of the preposition ‘in’ with respect to the morphological details of the container. We propose that the active semantic effect of ‘in’ is to trigger routines that transform a scene in the following manner: the container LM (‘bowl/cage’) is closed by adding virtual parts to become a continuous spherical or blob-like domain, while the contained object TR (‘fruit/bird’) expands by some contour diffusion process, irrespective of its detailed shape, until it collides with the boundaries of the closed LM. In fact, these two phases of the ‘in’ routine can often run in parallel: both TR and LM expand simultaneously until they reach each other’s boundaries (Fig. 1). The expansion of LM naturally creates its own closure while providing an obstacle to the expansion of TR. Therefore, we propose the following active semantic definition of the prototypical ‘containment’ schema (‘in₁–in₂’ and closely related variants): a domain TR lies ‘in’ a domain LM if an isotropic expansion of TR is stopped by the boundaries of LM’s own expansion.

A directly measurable consequence of this definition is the fact that no TR-induced activity will be detected on the borders of the visual field. This can be easily implemented in a 2-D square lattice, three-state CA (Fig. 1). The final boundary line between the domains is approximately equal to the skeleton of the complementary space between the

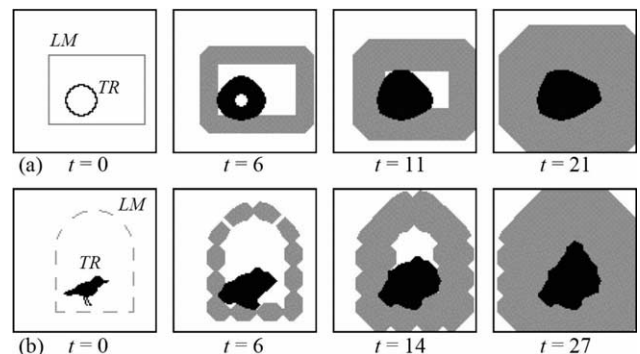


Fig. 1. Detection of a Proto-‘in’ Schema in a 2-D CA. We use a 100×100 , 3-state lattice: 1 for TR (black), 2 for LM (gray) and 0 for the ambient space (white). A simple diffusion rule requires that a 0 cell adopt the state of any non-0 nearest neighbor, if it exists. Starting with any two initial TR and LM domains and stopping when the total TR activity is constant, the lattice quickly converges to an attractor characterized by two contiguous domains of 1’s and 2’s. TR’s expansion is blocked by LM’s expansion from reaching the borders despite discontinuities in LM. (a) ‘The ball in the box’ converges in 21 steps. (b) ‘The bird in the cage,’ in 27 steps.

objects, also called *skeleton by influence zones* or SKIZ. In the case of the containment schema, the SKIZ surrounds TR and no cell on the borders is in TR's state 1. The sheer absence of TR activity on the borders is a robust feature that contributes to categorize the scene as 'in'. It illustrates a typical morphodynamical routine at the basis of our 'perceptual-semantic' classifier (Doursat & Petitot, 1997).

2.4. Translation invariance

Another example, 'the lamp is above the table', can be processed in the same way: the simultaneous expansion of TR and LM creates a roughly horizontal SKIZ line in the center of the image (Fig. 2). This time, the key classification feature is the absence of TR activity on the *bottom* border of the field, in conjunction with high TR activity on the top border and partial presence on the sides. This property is conserved by translation of TR in a broad region of space. Prototypical 'above' is one example of the generic 'partition' schema that includes elements such as 'below', 'beside', 'behind' and 'in front of' (viewed in 3-D).

Although rather simple, the double expansion process that we propose for the 'containment' and 'partition' schemas is nonetheless crucial and leads us to introduce the following general principles of active semantic morphodynamics: (i) objects have a tendency to occupy the whole space; (ii) objects are obstacles to each other's expansion. Through the action of structuring routines, the common space shared by the objects is divided into influence zones. Image elements cooperate to propagate activity across the field and inhibit activity from other sources.

2.5. Shape invariance → singularities

Talmy clearly shows that salient aspects of shapes are simply not taken into account by grammatical elements. In the sentences 'I zigzagged/circled/dashed through the woods' (Talmy, 2000, p. 27), the trajectory of TR, whether made of segments, curves or one straight line, ultimately joins two opposite (and possibly virtual) sides of an extended domain LM. In our active semantic view, the element 'through' creates a sharp schematization of TR in two steps: (1) tubification—a limited outward expansion convexifying the object—followed by

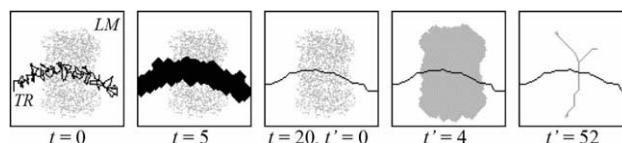


Fig. 3. Detection of a Proto-'through' Schema in a 2-D CA (same as Fig. 1). A tubification followed by a skeletonization is performed separately on TR (20 steps) and LM (52 steps), revealing an intersection point.

(2) skeletonization—the reverse inward expansion eroding the object down to its central symmetry axis. At the same time, it does the same with LM: the texture is first erased, then the domain is skeletonized (Fig. 3). The grammatical perspective enforced by 'through' thus reduces the detailed trajectory of TR or texture of LM to mere fluctuations around their medial axes. Finally, the characteristic feature of the 'through' schema lies at the intersection of the two skeletons.

Fig. 4 illustrates another scenario, TR 'out of' LM ('the ball out of the box'). In this case, the T-junction in the SKIZ between TR and LM accelerates away from TR and eventually disappears as TR exits the interior of LM. Again, this is a very robust bifurcation phenomenon, independent of the detailed shapes or trajectories of the objects.

In these two examples, 'through' and 'out of', the spatial relation between the objects is inferred from virtual *singularities* in the boundaries of the objects' territories. These singularities and their dynamical evolution are important clues that constitute the characteristic 'signature' of the spatial relationship (Petitot, 1995). Transformation routines considerably reduce the dimensionality of the input space, literally 'boiling down' the input images to a few critical features. A key idea is that singularities encode a lot of the image's geometrical information in an extremely compact and localized manner.

In summary, after the morphodynamical routines have transformed the scene according to the expansion principles (i) and (ii), several types of characteristic features can be detected to contribute to the final categorization: (iii) presence or absence of activity on the borders; (iv) intersection of skeletons; (v) singularities in the SKIZ boundary line. Our active semantic approach proposes a link between the 'things-relations-events' trilogy of cognitive

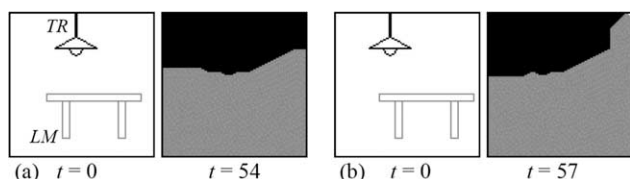


Fig. 2. Detection of a Proto-'above' Schema in a 2-D CA. We use the same lattice as Fig. 1. (a) LM's expansion prevents TR's expansion from reaching the bottom border. (b) Same effect when LM and TR are not vertically aligned, so that part of TR is directly facing the bottom border.

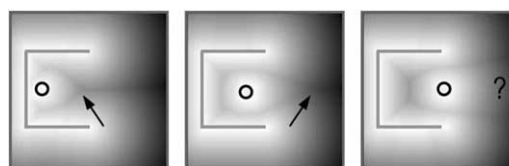


Fig. 4. Detection of a Proto-'out of' Schema in a 2-D CA. Each image is equivalent to the last step of Figs. 1 and 2, with TR and LM pasted on top. The influence domains of the ball and box are displayed in gray levels reflecting the distance to the objects (the farther, the darker). The sequence of images, in which TR moves out of LM, reveals a dynamical bifurcation (or 'phase transition', or 'catastrophe') as the singularity formed by the T-junction in the SKIZ boundary line disappears (shown by arrows).

grammar (Langacker, 1987) and the ‘domains-singularities-bifurcations’ trilogy of morphodynamical visual processing (Petitot, 1995). We suggest that the brain might rely on dynamical activity patterns of this kind to perform invariant spatial categorization.

3. What is ‘cognitive topology’?

The core invariance of spatial meaning reviewed above is sometimes referred to as the ‘linguistic form of topology’ or ‘cognitive topology’. Like mathematical topology (MT), language topology (LT) is magnitude and shape-neutral, yet in a way more abstract and less abstract at the same time (Talmy, 2000, p. 30). In some areas, LT has a greater power of generalization than MT. For example, ‘in’ closes the missing parts of ‘cage’ and ‘bowl’ to make them like a ‘box’. In other cases, however, LT preserves metric ratios and limits distortions in a stricter sense than MT. For example, the lexical elements ‘cup’ and ‘plate’ have distinct uses, although a plate is a flatter cup. The grammatical element ‘across’ also exhibits subtle metric constraints in its applicability: ‘he swam across the pool lane’ implies that the swimmer’s trajectory crosses the rectangle of water parallel to its short side, not its long side. Thus, because of these discrepancies LT has actually little to do with MT.

In geometry, one can identify six main levels of increasing structural constraints: sets (points), topological (‘glued’ points), differentiable (‘smooth’ objects), conformal (angle-preserving), metric (distance, e.g. Riemannian) and linear (vectors, e.g. Euclidean) spaces. We think that LT does not correspond to any of these levels or intermediate levels. In reality, the active conception of semantics presented above leads us to consider *interlevel transformations*, i.e. processes starting at one level and going up or down the hierarchy. For example, a convexification routine makes objects smoother and ‘rounder’: it discards most of the objects’ metric properties, therefore seems to go down, towards the ‘soft’ topological levels. Yet, at the same time, the convex classes it creates are more ‘rigid’ than the original objects because, precisely, mappings at their level must preserve convexity. Therefore, convex classes also belong to the higher metric levels. This illustrates the subtle interplay between impoverishment of details and rigidification of structures. Active semantics is neutral toward shape or continuity but it transforms objects into *prototypes*, which by nature are more metrically constrained. One could say that schematization and categorization replace ‘soft complex’ forms (e.g. intricate amoeboid blobs) with ‘rigid simple’ ones (e.g. spheres). Here, lies the puzzling apparent paradox of LT.

In summary, what is called ‘topology’ in Gestalt semantics and cognitive grammar actually corresponds to a family of *morphological operators*, as reviewed in Section 2. These operators are similar to the basic

operations of *dilation* and *erosion* defined in ‘the mathematical morphology’ (MM) of Serra (1982). Dilation and erosion can be combined to obtain the *closure* (filling small gaps), *opening* (pruning thin segments) or skeleton of a shape. Thus, MM is better suited than MT as a ‘toolbox’ for an active-semantic LT.

4. Waves in spiking neural networks

4.1. Dynamic pattern formation in excitable media

Elaborating upon this first morphodynamical model, we now establish a parallel with neural modeling. Our main hypothesis is that the transition from analog to symbolic representations of space might be neurally implemented by *traveling waves in a large-scale network of coupled spiking units*, via the expansion processes discussed above (see Fig. 6 for a preview). There is a vast cross-disciplinary literature, revived by Winfree (1980), on the emergence of ordered patterns in *excitable media* and coupled oscillators. Traveling or kinematic waves are a frequent phenomenon in nonlinear chemical reactions or multicellular structures (e.g. Swinney & Krinsky, 1991), such as slime mold aggregation, heart tissue activity, or embryonic pattern formation. Across various dynamics and architectures, these systems have in common the ability to reach a critical state from which they can rapidly bifurcate between randomness or chaos and ordered activity. To this extent, they can be compared to ‘sensitive plates’, as certain external patterns of initial conditions (chemical concentrations, food, electrical stimuli) can quickly trigger internal patterns of collective response from the units.

We explore the same idea in the case of an input image impinging on a layer of neurons and draw a link between the produced response and categorical invariance. In the framework proposed here, a visual input is classified by the qualitative behavior of the system, i.e. the presence or absence of certain singularities in the response patterns.

4.2. Spatiotemporal patterns in neural networks

During the past two decades, a growing number of neurophysiological recordings have revealed precise and reproducible temporal correlations among neural signals and linked them with behavior (Abeles, 1982; Bialek, Rieke, de Ruyter van Steveninck, & Warland, 1991; O’Keefe & Recce, 1993). *Temporal coding* (von der Malsburg, 1981) is now recognized as a major mode of neural transmission, together with average rate coding. In particular, quick onsets of transitory phase locking have been shown to play a role in the communication among cortical neurons engaged in a perceptual task (Gray, König, Engel, & Singer, 1989).

While most experiments and models involving neural synchronization were based on zero-phase locking among

coupled oscillators (e.g., Campbell & Wang, 1996; König & Schillen, 1991), *delayed* correlations have also been observed (Abeles, 1982, 1991). These nonzero-phase locking modes of activity correspond to reproducible rhythms, or waves, and could be supported by connection structures called *synfire chains* (Abeles, 1982; Ikegaya et al., 2004). Bienenstock (1995) construes synfire chains as the physical basis of elementary ‘building blocks’ that compose complex cognitive representations: synfire patterns exhibit *compositional* properties (Abeles, Hayon, & Lehmann, 2004; Bienenstock, 1996), as two waves propagating on two chains in parallel can lock and merge into one single wave by growing cross-connections between the chains (in a ‘zipper’ fashion). In this theory, spatiotemporal patterns in long synfire chains would thus be analogous to proteins that fold and bind through conformational interactions.

The present work construes wave patterns differently: we look at their emergence on regular 2-D lattices of coupled oscillators to implement the expansion dynamics of our morphodynamical spatial transformations. Compared to the traditional blocks of synchronization, i.e. the phase plateaus often used in segmentation models, we are also more interested in traveling waves, i.e. phase *gradients*.

4.3. Wave propagation and morphodynamical routines

A possible neural implementation of the morphodynamical engine at the core of our model relies on a network of individual spiking neurons, or local groups of spiking neurons (e.g. columns), arranged in a 2-D lattice similar to topographic cortical areas, with nearest-neighbor coupling. Each unit obeys a system of differential equations that exhibit regular oscillatory behavior in a certain region of parameter space. Various combinations of oscillatory dynamics (relaxation, stochastic, reaction-diffusion, pulse-coupled, etc.) and parameters (frequency distribution, coupling strength, etc.) are able to produce waveform activity, however, it is beyond the scope of the present work to discuss their respective merits. We want here to point out the generality of the wave propagation phenomenon, rather than its dependence on a specific model.

For practical purposes, we use Bonhoeffer-van der Pol (BvP) relaxation oscillators (FitzHugh, 1961). Each unit i is located on a lattice point x_i and described by a pair of variables (u_i, v_i) . Unit i is locally coupled to neighbor units j within a small radius r and may also receive an input I_i :

$$\begin{cases} \dot{u}_i = c(u_i - u_i^3/3 + v_i + z) + \eta + k \sum_j (u_j - u_i) + I_i \\ \dot{v}_i = (a - u_i - bv_i)/c + \eta \end{cases} \quad (1)$$

where η is a Gaussian noise, $\|x_i - x_j\| < r$ and $I_i = 0$ or a constant I . Parameters are tuned as in Fig. 5(a), so that individual units are close to a bifurcation in phase space between a fixed point and a limit cycle, i.e. one spike

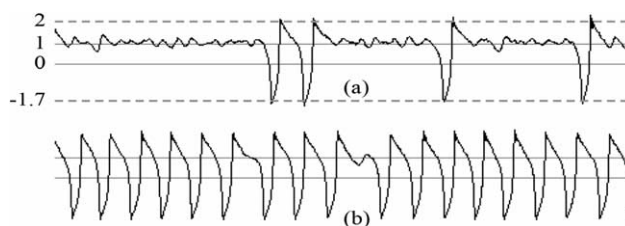


Fig. 5. Typical Firing Modes of a Stochastic BvP Relaxation Oscillator. Plots show $u(t)$, solution of Eq. (1) with $a = .7, b = .8, c = 3, \eta > 0, k = 0, I = 0$. (a) Sparse stochastic firing at $z = -.2$ (spikes are upside-down). (b) Quasi-periodic firing at $z = -.4$. At critical value $z_c = -.3465$, without noise, there is a bifurcation from a stable fixed point $u \approx 1$ to a limit cycle.

emission. They are *excitable* in the sense that a small stimulus causes them to jump out of the fixed point and orbit the limit cycle, during which they cannot be disturbed again.

Fig. 6(b) shows waves of excitation in a network of coupled BvP units created by the schematic scene ‘a small blob above a large blob’. Block impulses of spikes trigger wave fronts of activity that propagate away from the object contours and collide at the SKIZ boundary between the objects. These fronts are ‘grass-fire’ traveling waves, i.e. single-spike bands followed by refraction and reproducing only as long as the input is applied. Under the nonlinear dynamics, waves annihilate when they meet, instead of adding up. Fig. 6(a) shows the same influence zones obtained by mutual expansion in a CA, as seen in Section 2 (Fig. 2). Again, there is convincing perceptual and neural evidence for the significant role played by this virtual SKIZ structure and propagation in vision (Kimia, 2003).

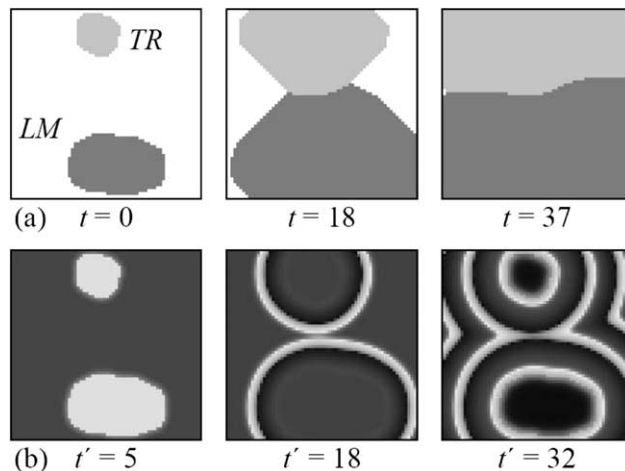


Fig. 6. Realizing Morphodynamical Routines in a Spiking Neural Network. (a) SKIZ obtained by diffusion in a 64×64 3-state CA, as in Fig. 2. (b) Same SKIZ obtained by traveling waves on a 64×64 lattice of coupled BvP oscillators in the regime of Fig. 5(a) with $\eta = 0$, connectivity radius $r = 2.3$ and coupling strength $k = .04$. Activity u is shown in gray levels, brighter for lower values, i.e. spikes $u < 0$. Starting with uniform resting potentials $u \approx 1$ (or weak stochastic firing with noise $\eta > 0$), an input image is continuously applied with amplitude $I = -.44$ in both TR and LM domains. This amounts to shifting z to a subcritical value $z = -.3467 < z_c$, thus throwing the BvP oscillators into periodic firing mode (Fig. 5(b)). This in turn creates traveling waves in the rest of the network.

5. Two wave-based categorization models

We now show how wave dynamics can support the categorization of spatial schemas by proposing two models based on the principles discussed in Section 2. First, waves implement the expansion-based transformations stated in principles (i)–(ii), then the detection of global activity or singularities created by the wave collisions is based on principles (iii)–(v). One wave model implements the border detection principle (iii) used in the ‘containment’ (Fig. 1) and ‘partition’ (Fig. 2) schemas. The second model focuses on the SKIZ singularities and ‘signature’ detection principle (v), which can be used as a complement or alternative to border detection. In this case, we illustrate SKIZ detection with the same ‘above’ schema as in border detection.

5.1. Border detection with cross-coupled lattices

Detecting the presence or absence of TR activity on the borders of the image, as for ‘in’ or ‘above’, is not possible in the single lattice of Fig. 6(b) because the waves triggered by TR and LM cannot be distinguished from each other. However, as discussed earlier in Section 4, a number of models have shown that lattices of coupled oscillators can also carry out *segmentation from contiguity* by exploiting a simpler form of temporal organization in the lattice: zero-phase synchronization or ‘temporal tagging’. We take here these results as the starting point of our simulation and assume that the original input layer is now split into two distinct sublayers, each holding one component of the scene. Thus, after a preliminary segmentation phase (not presented here), TR and LM are forwarded to layers L_{TR} and L_{LM} , where they generate wave fronts separately (Fig. 7). Mutual wave interference and collision is then recreated by *cross-coupling* the layers: unit i in layer L_{TR} is not only connected to units j in L_{TR} inside a neighborhood of radius r , but also to units j' in L_{LM} inside a neighborhood of radius r' .

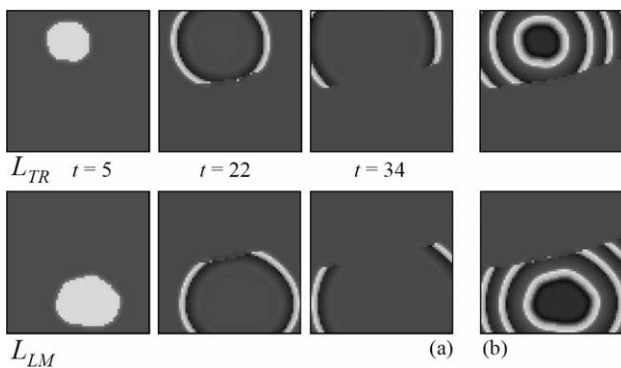


Fig. 7. Detection of the ‘above’ Schema by Mutually Inhibiting Waves. Two 64×64 lattices of BvP units, L_{TR} and L_{LM} , are internally coupled and cross-coupled according to Eq. (2) and its symmetrical version, with $r = r' = 2.3$, $\eta = 0$, $k = .03$, $k' = -.03$. (a) Single wave fronts obtained by injecting a pulse input $I = -.44$ in TR and LM for $0 \leq t < 2$ (10 time steps $dt = .2$). (b) Multiple wave fronts obtained by applying the same input amplitude indefinitely. In both cases, no spike reaches the bottom of L_{TR} .

The modified dynamics is therefore:

$$\begin{aligned} \dot{u}_i^{TR} = & F(u_i^{TR}) + k \sum_j \left(u_j^{TR} - u_i^{TR} \right) \\ & + k' \sum_{j'} \left(u_{j'}^{LM} - u_i^{TR} \right) + I_i^{TR} \end{aligned} \quad (2)$$

where $F(u)$ is the right hand side of Eq. (1) without k and I terms. A symmetrical relation holds for u_i^{LM} , swapping TR and LM. Variables v_i are not coupled and obey Eq. (1) as before. The net effect is shown in Fig. 7: the spiking wave fronts created by TR are cancelled by LM’s wave fronts and never reach the bottom border of L_{TR} , while hitting the top and partly the sides. This could be easily detected by external cells receiving afferents from the border units and linked to an ‘above’ response (not presented here). Again, the invisible collision boundary line is the SKIZ, which we now examine more closely in the next network model.

5.2. SKIZ signature detection with complex cells

Border activity provides a simple categorization mechanism but is generally not sufficient. Alone, it does not allow to distinguish among similar but nonoverlapping protoschemas, such as the ones in Fig. 8. This is where the properties of the SKIZ can help. For example, the concave or convex shape of the SKIZ is able to separate (b) and (c). For (a) and (d), one should also take into account the *flow velocity* along the SKIZ (Blum, 1973). Indeed, the dynamics of coupled spiking units (Fig. 6(b)) is richer than the morphological model (Fig. 6(a)) because it contains specific patterns of activity that are absent from a static geometric line. In particular, the wave fronts highlight a secondary flow of propagation *along* the SKIZ line, which travels away from the focal shock point with decreasing speed on either branch. The focal point (where the bright band is at its thinnest in Fig. 6(b), $t' = 32$) is the closest point to both objects and constitutes a *local optimum* along the SKIZ. While a great variety of object pairs produce the same static SKIZ, the speed and direction of flow along the SKIZ vary with the objects’ relative curvature and proximity to each other. For example, a vertical SKIZ segment flows outward between brackets

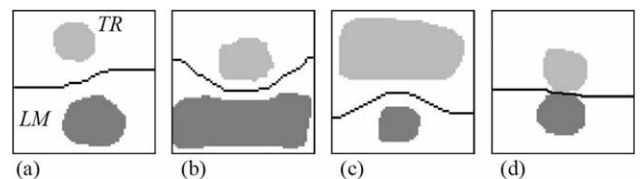


Fig. 8. Four Prototypical Horizontal Partition Schemas. In each case, TR (light gray) and LM (dark gray) are displayed with their SKIZ line (black). (a–c) English ‘above’. (b) Mixtec ‘siki’: LM is horizontally elongated (Regier, 1996). (c) French ‘par-dessus’: TR is horizontally elongated and covers LM. (d) English ‘on top of’: TR is in contact with LM. A simple border-based detection would give the same output in all cases.

facing their convex sides \rfloor (, whereas it flows inward between reversed brackets \lrcorner). This refined information is revealed by wave propagation.

In order to detect the focal points and flow characteristics (speed, direction) of the SKIZ, we propose in this model additional layers of detector neurons similar to the so-called ‘complex cells’ of the visual system. These cells receive afferent contacts from local fields in the input layer and respond to segments of moving wave bands, with selectivity to their orientation and direction. More precisely, the spiking neural network presented in Fig. 9 is a three-tier architecture comprising: (a) two input layers, (b) two middle layers of orientation and direction-selective ‘D’ cells, and (c) four top layers of coincidence ‘C’ cells responding to specific pairwise combinations of D cells. (These are not literally cortical layers but could correspond to functionally distinct cortical areas.) As in the previous model, TR and LM are separated on two independent layers (Fig. 9(a)). In this particular setup, however, there is no cross-coupling between layers and the waves created by TR and LM do not actually interfere or collide. Rather, the regions where wave fronts coincide ‘vertically’ (viewing layers L_{TR} and L_{LM} superimposed) are captured by higher feature cells in two

direction-selective layers, D_{TR} and D_{LM} (Fig. 9(b)), and four coincidence-detection layers, $C_{1..4}$ (Fig. 9(c)). Layer D_{TR} receives afferents only from L_{TR} , and D_{LM} only from L_{LM} . Layers C_i are connected to both D_{TR} and D_{LM} through split receptive fields: half of the afferent connections of a C cell originate from a half-disc in D_{TR} and the other half from its complementary half-disc in D_{LM} .

In the intermediate D layers, each point contains a family of cells or ‘jet’ similar to multiscale Gabor filters. Viewing the traveling waves in layers L as moving bars, each D cell is sensitive to a combination of bar width λ , speed s , orientation θ and direction of movement ϕ . In the simple wave dynamics of the L layers, λ and s are approximately uniform. Therefore, a jet of D cells is in fact single-scale and indexed by one parameter $\theta=0\dots2\pi$, with the convention that $\phi=\theta+\pi/2$. Typically, 8 cells with orientations $\theta=k\pi/4$ are sufficient. Each sublayer $D_{TR}(\theta)$ thus detects a portion of the traveling wave in L_{TR} (same with LM). Realistic neurobiological architectures generally implement direction-selectivity using inhibitory cells and transmission delays. In our simplified model, a $D(\theta)$ cell is a ‘cylindrical’ filter, i.e. a temporal stack of discs containing a moving bar at angle θ (Fig. 9(b), center of D layers): it sums potentials from afferent L-layer spikes spatially and temporally, and fires itself a spike above some threshold.

Among the four top layers (Fig. 9(c)), C_1 detects converging parallel wave fronts, C_2 detects diverging parallel wave fronts, and C_3 and C_4 detect crossing perpendicular wave fronts. Like the D layers, each C_i layer is subdivided into 8 orientation sublayers $C_i(\theta)$. Each cell in $C_1(\theta)$ is connected to a half-disc neighborhood in $D_{TR}(\theta)$ and the complementary half-disc in $D_{LM}(\theta-\pi)$, where the half-disc separation is at angle θ . The net output of this hierarchical arrangement is a *signature* of coincidence detection features providing a *very sparse coding* of the original spatial scene. The input ‘above’ scene is eventually reduced to a handful of active cells in a single orientation sublayer $C_i(\theta)$ for each C_i : $C_1(\pi)$, $C_2(0)$, $C_3(3\pi/4)$ and $C_4(-\pi/4)$ (Fig. 9(c)).

In summary, the active cells in C_1 and C_2 reveal the focal point of the SKIZ, which is the primary information about the scene, while C_3 and C_4 reveal the outward flow on the SKIZ branches, which can be used to distinguish among similar but nonequivalent concepts. This sparse SKIZ signature is at the same time characteristic of the spatial relationship and largely insensitive to shape details. For example: ‘below’ yields $C_1(0)$ and $C_2(\pi)$; ‘on top of’ (Fig. 8(d)) yields $C_2(0)$ like ‘above’ but no C_1 activity because TR and LM are contiguous (wave fronts can only separate at the contact point, not join); ‘par-dessus’ with a convex SKIZ facing up (Fig. 8(c)) yields $C_3(\pi)$ and $C_4(-\pi/2)$, etc. Note that the actual regions of $C_i(\theta)$ where cells are active (e.g. the location of the SKIZ branches in the south–west and south–east quadrants of Fig. 8(c)) are sensitive to translation and therefore are not good invariant features.

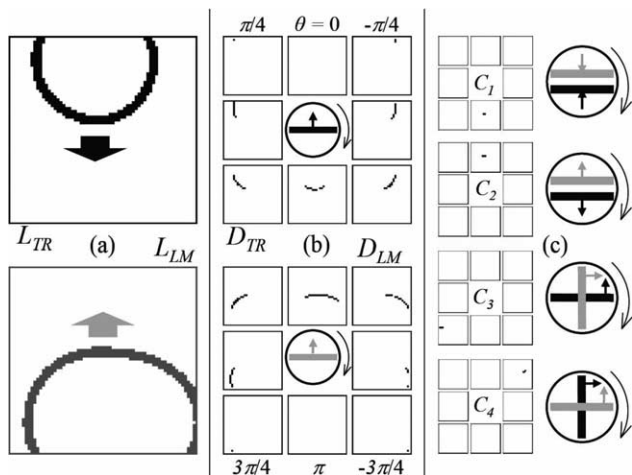


Fig. 9. SKIZ Detection in a Three-Tier Spiking Neural Network Architecture. (a) Input layers show single traveling waves as in Fig. 7(a), except that there is no cross-coupling and potentials are thresholded to retain only the spikes $u < 0$. (b) Orientation and direction-selective cells. Each D layer is shown as 8 sublayers (smaller squares) of cells selective to orientation $\theta = k\pi/4$, with $k = 8 \dots 1$ from the top, clockwise. (c) Pairwise coincidence cells. Each cell in $C_1(\theta)$ is connected to a half-disc neighborhood in $D_{TR}(\theta)$ and its complementary half-disc in $D_{LM}(\theta - \pi)$, rotated at angle θ . For example: a cell in $C_1(0)$ (illustrated by the icon on the right of C_1) receives afferents from a horizontal bar of $D_{TR}(0)$ cells in the lower half of its receptive field, and a bar of $D_{LM}(\pi)$ cells in the upper half. Same for $C_2(0)$, swapping TR/LM and upper/lower. Similarly, $C_3(\theta)$ cells are half-connected to $D_{TR}(\theta)$ and to $D_{LM}(\theta - \pi/2)$, with orthogonal bars (swapping TR/LM for C_4). The net output is sparse activity confined to $C_1(\pi)$, $C_2(0)$, $C_3(3\pi/4)$ and $C_4(-\pi/4)$. We use only global C_i activity: in reality, C_1 cells fire first when the two wave fronts meet, then C_2 cells when they separate again, and finally C_3 and C_4 cells in rapid succession when the arms cross. This precise rhythm of C_i spikes could also be exploited in a finer model.

6. Discussion

We have proposed a dynamical approach to cognitive linguistics drawing from morphological CA and spiking neural networks. We suggest that spatial semantic categorization can be supported by an expansion-based dynamics, such as activity diffusion or wave propagation. Admittedly, the few results we have presented here are not particularly surprising given the relative simplicity and artificiality of the models. However, we hope that this preliminary study will be a starting point for more efficient or plausible network architectures exploring the interface between high-level vision and symbolic knowledge.

6.1. Future work

- (1) *Wave Dynamics and Scene Database*: we want to conduct a more systematic investigation of the morphodynamical routines and their link with proto-semantic classes. Similarly to Regier (1996), a database of schematic image/label pairs representing a broad cross-linguistic variety of spatial elements could be used to assess the level of invariance of the singularities and their robustness to noise.
- (2) *Real Images and Low-Level Vision*: currently, our primary material consists of presegmented schematic images. It could be extended to real-world examples by using low-level image processing techniques based on edge contiguity and texture. Segmentation models such as nonlinear diffusion (Whitaker, 1993), variational boundary/domain optimization (Mumford & Shah, 1989) or temporal phase tagging (König & Schillen, 1991) have proven that shapes can be separated from the background in a bottom-up way without prior knowledge, if the scene is not too cluttered.
- (3) *Learning the Semantics from the Protosemantics*: semantic classes are intrinsically fuzzy: as TR moves around LM and their SKIZ rotates, when is TR no longer ‘above’ but ‘beside’ or ‘below’ LM? Different languages also divide space differently: for example, ‘on’ corresponds in German either as ‘auf’ (top contact) or as ‘an’ (side contact). Intra- and cross-linguistic boundaries could be learned using known image/response pairs. Our morphodynamical routines already considerably reduce the dimensionality of the input space by mapping images to a few singularities. In a final step, the same universal pool of protosemantic features could be combined in various ways to form full-fledged semantic classes using statistical estimation methods. What is critical, however, is how well the elementary routines map the examples into optimally separable clusters (Geman, Bienenstock, & Doursat, 1992). Applying learning methods directly to the original image space or to irrelevant features would evidently be fruitless.
- (4) *Verb Processes and Bifurcation Events*: another important challenge are the temporal processes and events

of verbal scenarios. The singularities created by fast wave activity can themselves evolve on a slower timescale (Fig. 4). Short animated clips of moving TR’s and LM’s could be categorized into archetypal verbs, e.g. ‘give’ or ‘push’, based on an important family of psychological experiments about the perception of causality and animacy. Landmark studies (see Scholl & Tremoulet, 2000) have shown that movies involving simple geometrical figures were spontaneously interpreted by human subjects as intentional actions. For example, a few triangles and circles moving around a square strongly tend to elicit verbal statements such as ‘chase’, ‘hide’, ‘attack’, ‘protect’, etc.

- (5) *Complex Scenes*: after treating single schemas separately, we also want to show how multiple schemas can be evoked simultaneously and assembled to form complex scenarios. This addresses the *compositionality* of semantic concepts, or ‘binding problem’ (von der Malsburg, 1981). Our ultimate goal is to explain mental imagery in terms of structured compositions of morphodynamical routines.

6.2. Original points of this proposal

- (1) *Bringing Large-Scale Dynamical Systems to Cognitive Linguistics*: despite their deep insights into the conceptual and spatial organization of language, cognitive grammars still lack mathematical and computational foundations. Our project is among few attempts to import spatiotemporal connectionist models into linguistics and spatial cognition. Other authors (e.g. Regier, 1996; Shastri & Ajjanagadde, 1993) have pursued the same objectives, but use small ‘hybrid’ artificial neural networks, where nodes already carry geometrical or symbolic features. We work at the fine-grained level of numerous spatially arranged units.
- (2) *Addressing Semantics in CA and Neural Networks*: conversely, our work is also an original proposal to apply large-scale lattices of cellular automata or neurons to high-level semantic feature extraction. These bottom-up systems are usually exploited for low-level image processing or visual cortical modeling, or both—e.g. Pulse-Coupled Neural Networks (Johnson, 1994) or Cellular Neural Networks (Chua & Roska, 1998). Shock graphs and medial axes are also used in computer vision models of object recognition (Siddiqi, Shokoufandeh, Dickinson, & Zucker, 1999; Zhu & Yuille, 1996), but with the concern to preserve and match object shapes, not erase them. Adamatzky (2002) also envisions collision-based wave dynamics in excitable media, but as a mechanism of universal computing based on logic gates.
- (3) *Advocating Pattern Formation in Neural Modeling*: self-organized, emergent processes of pattern formation or *morphogenesis* are ubiquitous in nature

(stripes, spots, branches, etc.). As a complex system, the brain produces ‘cognitive forms’, too, but instead of spatial arrangements of molecules or cells, these forms are made of *spatiotemporal patterns* of neural activities (synchronization, correlations, waves, etc.). In contrast to other biological domains, however, pattern formation in large-scale neural networks has attracted only few authors (e.g. Ermentrout, 1998). This is probably because precise rhythms involving a large number of neurons are still experimentally difficult to detect, hence not yet proven to play a central role.

- (4) *Suggesting Wave Dynamics in Neural Organization*: indeed, fast rhythmic activity on the 1-ms timescale, in random (Bienenstock, 1995), regular (Milton, Chu, & Cowan 1993) or small-world (Izhikevich, Gally, & Edelman, 2004) networks, has been much less explored than pure synchronization, when addressing segmentation or the ‘binding problem’ (see review in Roskies, 1999). Along with Bienenstock (1995, 1996), we contend that waves open a richer space of temporal coding suitable for general *mesoscopic* neural modeling, i.e. the intermediate level of organization between microscopic neural activities and macroscopic representations. At one end (AI), high-level formal models manipulate symbols and composition rules but do not address their fine-grain internal structure. At the other end (neural networks), low-level dynamical models study the self-organization of neural activity but their emergent objects (attractor cell assemblies or blobs) still lack structural complexity and compositionality (Fodor & Pylyshyn, 1988). Waves and other complex spatiotemporal patterns could provide the missing link bridging the gap between these two levels. Furthermore, our hypothesis is that this mesoscopic level corresponds to the central conceptual level postulated by cognitive linguistics.

Acknowledgements

R.D. is grateful to Philip Goodman for welcoming him in his UNR lab, and to J.P. for his warm encouragements.

References

- Abeles, M. (1982). *Local cortical circuits*. Berlin: Springer-Verlag.
- Abeles, M. (1991). *Corticonics*. Cambridge: Cambridge University Press.
- Abeles, M., Hayon, G., & Lehmann, D. (2004). Modeling compositionality by dynamic binding of synfire chains. *Journal of Computational Neuroscience*, 17, 179–201.
- Adamatzky, A. (Ed.). (2002). *Collision-based computing*. Berlin: Springer-Verlag.
- Bialek, W., Rieke, F., de Ruyter van Steveninck, R., & Warland, D. (1991). Reading a neural code. *Science*, 252, 1854–1857.
- Bienenstock, E. (1995). A model of neocortex. *Network*, 6, 179–224.
- Bienenstock, E. (1996). Composition. In A. Aertsen, & V. Braitenberg (Eds.), *Brain theory*. Elsevier.
- Blum, H. (1973). Biological shape and visual science. *Journal of Theoretical Biology*, 38, 205–287.
- Campbell, S., & Wang, D. (1996). Synchronization and desynchronization in a network of locally coupled Wilson-Cowan oscillators. *IEEE Transactions on Neural Networks*, 7, 541–554.
- Chua, L., & Roska, T. (1998). *Cellular neural networks and visual computing*. Cambridge, UK: Cambridge University Press.
- Doursat, R., & Petitot, J. (1997). Modèles dynamiques et linguistique cognitive: vers une sémantique morphologique active. In Actes de la 6ème École d’été de l’Association pour la recherche cognitive. CNRS.
- Doursat, R., & Petitot, J. (2005). Bridging the gap between vision and language: A morphodynamical model of spatial categories. Proceedings of the International Joint Conference on Neural Networks, July 31–August 4, 2005, Montréal, Canada.
- Ermentrout, B. (1998). Neural networks as spatio-temporal pattern-forming systems. *Reports on Progress in Physics*, 61, 353–430.
- FitzHugh, R. A. (1961). Impulses and physiological states in theoretical models of nerve membrane. *Biophysical Journal*, 1, 445–466.
- Fodor, J., & Pylyshyn, Z. (1988). Connectionism and cognitive architecture: a critical analysis. *Cognition*, 28(1/2), 3–71.
- Geman, S., Bienenstock, E., & Doursat, R. (1992). Neural networks and the bias/variance dilemma. *Neural Computation*, 4, 1–58.
- Gray, C. M., König, P., Engel, A. K., & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit intercolumnar synchronization which reflects global stimulus properties. *Nature*, 338, 334–337.
- Herskovits, A. (1986). *Language and spatial cognition: an interdisciplinary study of the prepositions in English*. Cambridge: Cambridge University Press.
- Ikegaya, Y., Aaron, G., Cossart, R., Aronov, D., Lampl, I., Ferster, D., & Yuste, R. (2004). Synfire chains and cortical songs: temporal modules of cortical activity. *Science*, 304, 559–564.
- Izhikevich, E. M., Gally, J. A., & Edelman, G. M. (2004). Spike-timing dynamics of neuronal groups. *Cerebral Cortex*, 14, 933–944.
- Jackendoff, R. (1983). *Semantics and cognition*. Cambridge, MA: The MIT Press.
- Johnson, J. L. (1994). Pulse-coupled neural nets: translation, rotation, scale, distortion and intensity signal invariance for images. *Applied Optics*, 33(26), 6239–6253.
- Kimia, B. B. (2003). On the role of medial geometry in human vision. *Journal of Physiology Paris*, 97(2/3), 155–190.
- König, P., & Schillen, T. B. (1991). Stimulus-dependent assembly formation of oscillatory responses. *Neural Computation*, 3, 155–166.
- Lakoff, G. (1987). *Women, fire, and dangerous things*. Chicago, IL: University of Chicago Press.
- Langacker, R. (1987). *Foundations of cognitive grammar*. Palo Alto, CA: Stanford University Press.
- Lee, T. S. (2003). Computations in the early visual cortex. In J. Petitot, & J. Lorenceau, *Journal of Physiology Paris* (vol. 97), 121–139 (2/3).
- Leyton, M. (1992). *Symmetry, causality, mind*. Cambridge, MA: The MIT Press.
- Mallat, S. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7), 674–693.
- Marr, D. (1982). *Vision*. San Francisco, CA: Freeman Publishers.
- Milton, J. G., Chu, P. H., & Cowan, J. D. (1993). Spiral waves in integrate-and-fire neural networks. In *Advances in neural information processing systems*, vol. 5. Los Altos, CA: Morgan Kaufmann (pp. 1001–1007).
- Mumford, D., & Shah, J. (1989). Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*, 42, 577–685.
- O’Keefe, J., & Recce, M. L. (1993). Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus*, 3, 317–330.
- Petitot, J. (1995). Morphodynamics and attractor syntax. In T. van Gelder, & R. Port (Eds.), *Mind as motion*. Cambridge, MA: The MIT Press.

- Regier, T. (1996). *The human semantic potential*. Cambridge, MA: The MIT Press.
- Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 104, 192–233.
- Roskies, A. L. (1999). The binding problem. *Neuron*, 24, 7–9.
- Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Science*, 4(8), 299–309.
- Serra, J. (1982). *Image analysis and mathematical morphology*. New York: Academic Press.
- Shastri, L., & Ajjanagadde, V. (1993). From simple associations to systematic reasoning. *Behavioral and Brain Sciences*, 16(3), 417–451.
- Siddiqi, K., Shokoufandeh, A., Dickinson, S. J., & Zucker, S. W. (1999). Shock graphs and shape matching. *International Journal of Computer Vision*, 35(1), 13–32.
- Swinney, H. L., & Krinsky, V. I. (Eds.). (1991). *Waves and patterns in chemical and biological media*. Cambridge, MA: The MIT Press.
- Talmy, L. (2000). *Toward a cognitive semantics Volume I: concept structuring systems*. Cambridge, MA: MIT Press (Original work published 1972–1996).
- von der Malsburg, C. (1981). *The correlation theory of brain function*. Internal Report 81-2 Max Planck Institute for Biophysical Chemistry.
- Whitaker, R. (1993). Geometry-limited diffusion in the characterization of geometric patches in images. *Computer Vision, Graphics, and Image Processing: Image Understanding*, 57(1), 111–120.
- Winfree, A. (1980). *The geometry of biological time*. Berlin: Springer-Verlag.
- Zhu, S. C., & Yuille, A. L. (1996). FORMS: a flexible object recognition and modeling system. *International Journal of Computer Vision*, 20(3), 187–212.